

# Machine Learning and Deep Learning: Object Detection Approaches

Parul Bhanarkar<sup>1</sup>

Department of Computer Science & Engineering, JIT, Nagpur, Maharashtra, India

## Abstract:

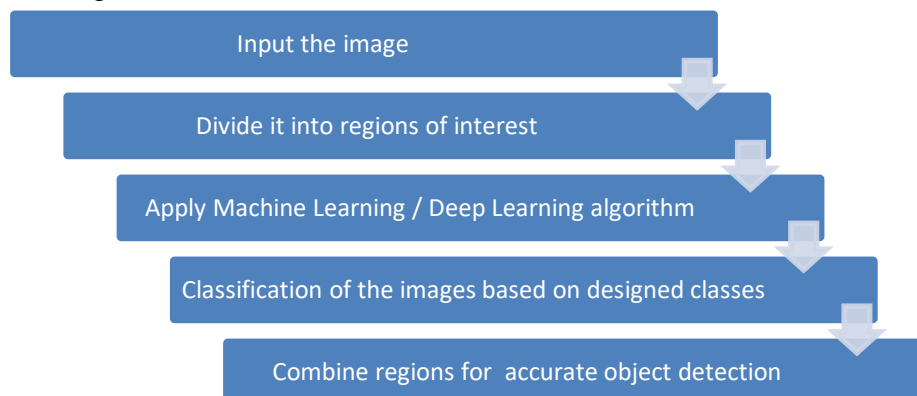
Object recognition strategies work on tracking down the presence of certifiable items in advanced pictures or recordings. The assurance of the items is finished through the distinguishing proof of classes utilized for the ID and recognition of articles. Different Computer vision assignments can be performed with the assistance of article discovery calculations like face acknowledgment, augmented reality and item following. Object recognition is effectively carried out utilizing profound learning and AI calculations. Object recognition can be called as the mix of article restriction and order. The precision of the removed elements is important to recognize the items in the picture accurately. The result produced by the item identification calculations is fundamentally a picture with a limited box with the name of the item recognized. The paper depicts a short survey on AI calculations for object location including R-CNN, Fast R-CNN,

Mask-RCNN, R-FCN, HOG, SPP-Net, SSD and YOLO. A definite conversation on profound learning approaches valuable for object discovery is likewise introduced.

**Keywords:** Machine Learning, Deep Learning, Object detection, object based classification, Evaluation parameters, Convolution Neural networks, classification accuracy.

## 1. Introduction:

As of late, PC vision undertakings performed by machines have impacted the market generally because of their exactness in playing out the given errands. The conventional techniques utilized for object recognition and the new strategies contrast in the exactness and speed of article identification. The fundamental subtasks associated with the article recognition process is depicted in the chart beneath:



**Figure. 1.1. Object Detection Process**

To find the locales of interest, area proposition strategies can be utilized like particular inquiry. Object restriction frames the essential errand in the discovery of items which comprises of tracking down the area of the article by making a limited box around the article inside the district of interest. The item restriction is considered as a relapse issue under AI. The IoU metric can be utilized to test the precision of the article limitation strategy. Characterization assumes a crucial part in performing object identification. Order is a kind of example acknowledgment task for instance sorting the sends into spam and no-spam. Not many famous AI calculations incorporate SVM, strategic relapse, closest neighbor and choice trees. Famous Deep Learning calculations which find application in object discovery incorporate CNN, LSTM, RNN, GANs, and so forth. The paper is separated into 3 segments. The primary segment depicts the item discovery strategies created under Machine Learning and Deep Learning. The following segment presents a similar investigation of the different machine and profound learning approaches examined in segment 1. The third segment portrays the examination and assessment of execution of few methodologies. A concise outline of the strategies is examined in the later segment.

## 2. Machine Learning and Deep Learning: Object Detection Methods

**D. Moshou, C. Bravo, R. Oberti, J. West, L. Bodria, A. McCartney, H. Ramon[1]**, the creators have worked over a discovery calculation applied over the yields to distinguish or identify the sickness in beginning phase. For the work, combination of hyperspectral reflection

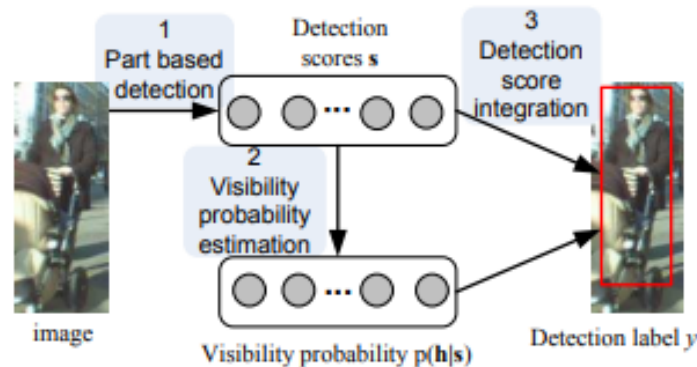
data between 450 to 900 nm including fluorescence imaging. The yellow rust illness of wheat is taken as a model for the framework testing. An imaging spectrograph was utilized to catch the pictures of the sound and contaminated plants under legitimate ecological circumstances. In the work, the negligible part of pixels which address the illness was set as sickness variable and is called as sore record (LI). For the assessment three wavebands were utilized to segregate the sickness on the plant. This technique depends on fluorescence is less precise. The creators then checked the precision of the consolidated strategy where information combination was performed utilizing a self getting sorted out map which was helpful in diminishing the general order mistake.

In the work, the technique in view of the ghostly reflectance estimations and other strategy in light of fluorescence enlistment, where the three waveband were chosen for the arrangement and illness discovery. In the fluorescence technique, for every estimation, two information pictures are taken with lighting and without lighting. However, just the picture at the fluorescence top was utilized for the result.

**W. Ouyang and X. Wang[2]**, the creators have taken care of over the issues with part based object identification strategy and created probabilistic passerby location technique. The disfigurements score are distinguished from the part based model and the visibilities are then demonstrated. This perceivability relationship has been mastered involving the discriminative profound model for the impediment taking care of. Different deformable models

which are being utilized for object discovery including the Generic finders which has normal execution; passerby existing information window has the high added score. The issue of impediment is extremely normal in input pictures in the event of swarmed scene. The location could find true success on the off chance

that we could incorporate the extra data of the body parts impeded in the picture. When the recognizable proof of the blocked body parts is finished, their impact can then be eliminated from the last score. The pedestrian detection with hidden occlusion variable approach is shown in the following figure:



**Figure. 2.3.1. Pedestrian detection using deep learning**

The work here relates with planning covering parts at numerous layers and later check the perceivability of the part at various layers. In this example, we can get the significant logical data for the specific assessment of the covering parts in the picture. The significant test in the work is the displaying of the visibilities of various parts and the relationship among them. This is then used to produce the consolidated consequences of the part finders as indicated by the valuation of the visibilities. The discriminative profound learning based model utilized here learns the relationship between the different body parts. This is additionally relevant in aspect decrease and acknowledgment.

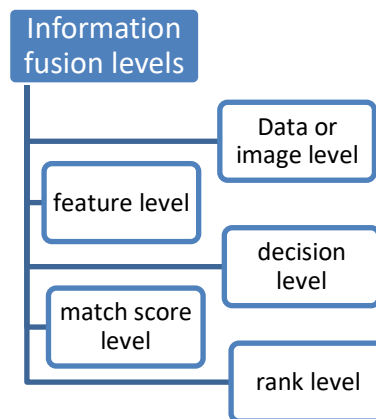
**Richa Singh, Mayank Vatsa, Afzel Noor [3]**, a technique for development of face acknowledgment has been examined in this work. The two levels progressive combination has been applied on the

pictures which are caught utilizing two different light ranges. The ranges utilized here are the apparent light and infrared light. The DWT based combination calculation was utilized for the two pictures at the picture level combination. The plentifulness and the stage highlights were used at the element combination level utilizing a 2D log polar Gabor Wavelett. The SVM has been subsequently applied to produce the combined combination set for further developing the face acknowledgment by and by. Certain provokes exist with regards to confront acknowledgment including the picture caught under noticeable light, looks, changes in the lighting conditions, impediments on the face, and so forth. These difficulties have been tended to in the work. The work plays out the preparation of the model utilizing single face picture. The elements separated utilizing the above strategy is liberated

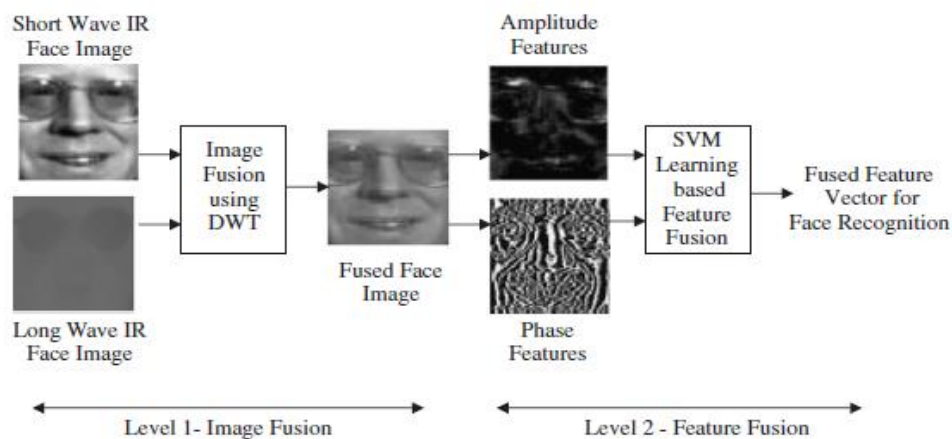
from the ecological imperatives like enlightenment and differentiation. For utilizing adequacy and face highlights in the given picture, the triangle face location calculation is utilized to distinguish the face and afterward it is addressed by picture  $I(x,y)$  in it be  $N \times N$  size. Then, at that point, it is changed into log polar structure after which it is switched over completely to  $I(r,b)$ .

The combination of picture level and component level is finished at the following stage to further develop the acknowledgment execution. The cycle is finished in five phases as displayed in figure beneath:

Since the conventional staggered brain networks were having the issues of creating various nearby minima arrangements, the 2v-SVM has been use to play out the component combination. This strategy generally finds a worldwide minima. The 2v-SVM is applied as the two class issue. Which thusly arranges the great component class and the awful element class. The figure beneath shows the element level combination:



**Figure. 2.3.2. Information fusion levels**



**Figure. 2.3.3. Two level hierarchical fusion**

The image level and feature level feature integration is done for enhancing the performance of the face recognition. The

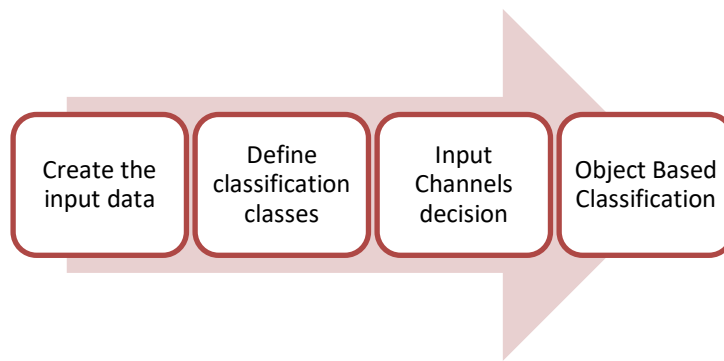
composite feature vector is then generated for further recognition. This algorithm is applicable for texture based recognition algorithm.

**Leitner, R., Biasio, M., De Arnold, T., Dinh, C.V., Loog, M., Duin, R.P.W[4]**, the creators have expressed that the already accessible multispectral picture obtaining framework had restricted limit as they were of less utilization towards endoscopy. This was because of the essential spatial examining and unrealistic long exchanging times in the fluid gem tunable channels. Here, the work addresses the Acousto-optical tunable channels which makes this changing times as low to catch the hyperspectral pictures and furthermore investigate the multispectral video information taking care of the spatial and fleeting goal. The procurement of the data of the hyperspectral pictures to be utilized for the hyperspectral imaging endoscopy for the recognition of carcinogenic tissues. The hyperspectral estimations for the biopsies are then taken, it is seen that the ghastly elements of the harmful districts are subject to the tissue. Here classifiers, for example, Support Vector Machines or Mixture of the Gaussian classifier can't effectively sum up the data. It has been seen that the preparation information determination conspire picking the very preparing tests for an equivalent test that accomplishes a superior exactness of expectation which utilizes QDC having further developed power and less possibilities overtraining. This is then

joined with the picture enlistment which helps eliminating the movement based procurement which thusly works with the increase of the video transfer with overlays. This could help in distinguishing and showing the destructive tissue areas.

The creators chipped away at the technique for video endoscopy that has the capacity of distinguishing the harmful and precancerous locales. This is finished by increase of the video transfer alongside an overlay featuring districts that need assessment. The contribution to such a framework is the spectroscopic estimations of biopsies having otherworldly finger impression giving significant phantom data to the determination. The creators have examined about the algorithmic blocks expected for playing out the multi otherworldly video endoscope which can distinguish the area. The multivariate order which is joined with the movement remedy calculation creates the video overlay valuable for disease tissue location. The outcomes have shown the expected utilization of the innovation for clinical imaging.

**Volker Walter [5]**, a methodology for the characterization of the gathering of pixels that address objects in a GIS data set. This strategy depends on a change location technique which depends on the item based order utilizing regulated greatest probability characterization. For the order, the n-layered highlight space is made from the proportions of multispectral groups. The preparation tests are taken from the GIS data set. In the work, the cycle followed is portrayed in the outline beneath:



**Figure. 2.3.4. Process of Object based classification over GIS data**

The remote detecting information was utilized in the work taken from the DPA framework, an optical airborne camera. A few item classes were distinguished including modern region, private area, etc. The classes were then partitioned into additional spaces. In the work, the creators have utilized all ghostly and as information channels. In the article based arrangement, every one of the pixels that are of that GIS object are essentially held together and are subsequently used to examine the otherworldly data about the item viable. The mean dim worth is determined for each channel for every one of the items. This is like smoothing the info information. Here the way of behaving of the article is same when contrasted with the ghastry way of behaving of the pixels. Further expansion of info channels are likewise conceivable if necessary. For acquiring the characterization results the info channel was traversed over the 16 layered include space. The methodology worked by bringing the items into the classes: water, woodland, settlement and Greenland. The work can be more refined if more qualities would be added and assessed. Here the articles were partitioned physically into

one of the three classes. The five star from the gathering contained every one of the items where the adjustment of the scene could be noted and update is to be finished. The subsequent one contains objects where GIS objects are accurately arranged or no can be tried. Here high goal information are expected to conclude whether the GIS information should be refreshed. The last class contains the articles that are mistaken.

**Ian Fasel, Bret Fortenberry, Javier Movellan[6]**, the creators have worked over the item finding and element extraction by a probabilistic model in view of specific ideal derivation calculations. Here the pictures are taken as montage of patches of erratic size, which might contain the article that should be distinguished or just the foundation. The creators have fostered a probability proportion model for something similar. The learning in these models are performed utilizing supporting techniques. The methodology was applied over countenances and eyes on erratic pictures. The work is finished on the acknowledgment of tracking down appearances and elements as Bayesian derivation issue. The model depends on finding how pictures are created and

involving the planned model making helpful derivations according to require. The probability proportion is determined in the deduction cycle over the given erratic picture fix utilizing the Gentle Boost calculation and the gauge between the proportion is determined by the probabilities around the fix created by the article class and the foundation class.

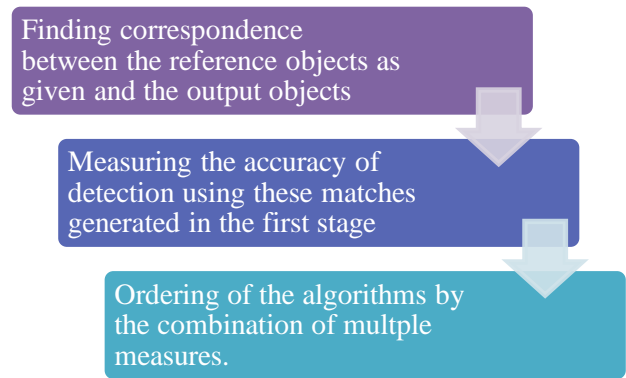
In the work referenced, video pictures continuously are thought about where each edge is dealt with autonomously. At the point when each edge is dealt with autonomously, it permits the framework to code the eye area and its way of behaving which might come all through the scene. The Gentle Boost calculation chooses wavelets from a huge pool and joins them to limit the blunder in the result. The wavelets for the most part comprises of Haar-like wavelets. The upside of utilizing it is that it produces quick result by taking the amount of pixels or four equivalent estimated and taking contrasts of these totals.

**Asaaki Shibata, Yuichiro Yasuda, Masahide Ito**[7], the creators have worked on the moving item identification on the pictures caught utilizing a functioning camera. In the strategy the creators have worked over the examination of optical streams for the discovery of items. The optical stream is taken care of with the assistance of picture data and camera movement. This optical stream data age is exactly finished by controlling the camera movements. One of the element of the dynamic moving camera is that it permits sufficient data connected with the climate more than the static cameras. The dynamic cameras have the likelihood to change its situation

and arrange itself. The creators have featured many recently included strategies including foundation deduction strategy, entomb outline contrast technique and a few different methods in light of ideal stream. The optical stream can be characterized as a vector which addresses the item speed in the given picture. The stream comprises of power and direction. These work as requirements for the boundaries. The optical stream is created utilizing the camera caught pictures. The recognition of moving items is a seriously difficult errand as the stream is created for the static and moving items when the camera is moving. This is misshaped stream since the items are moving in the dynamic camera. This bending has been utilized to separate between the moving articles. In this work, the power of the luminance present in the picture and the movement from the dynamic camera has been utilized to determine the optical stream at each point n time during the examinations.

**Bahadır Özdemir, Selim Aksoy, Sandra Eckert, Martino Pesaresi** [8], the creators have worked over the presentation boundaries assessment of the item recognition calculations in this work. The work specifies the correspondence between the reference and the result protests, the exactness score related to the item shape, limit, additionally the discontinuity mistakes and the positioning step utilizing the numerous signs of execution. With regards to checking the exhibitions of the item discovery calculations is to consider the pixel based characterization where the naming of picture pixels is utilized for the recognition. The mistake rates produced

from the disarray network is valuable for execution assessment. The pixel based assessment strategies are appropriate for applications like guide refreshing, change discovery, imperfection identification as well as deformity location helpful while recognizing the few pixels on the objects of interest. It makes the discovery of blunders and adjusts the calculations more straightforward. The different item based exhibition measures could be utilized to conquer the impediments engaged with the pixel based boundaries. This technique attempts to find the balanced correspondence between the reference object and the result object distinguished. There are different issues connected with the item based exhibition boundaries including the under location and the over recognition issues. The under location issues include the recognition of result object compares to more than one reference objects. The over recognition issues include the correspondence of more than one result object with one reference object. For this situation certain limit is permitted to be followed. The under and over location issues structure to be significant elements in the appraisal of the precision cycle where enormous number of items are to be thought of. The work here focuses on the one, many-one correspondence of the reference and result objects alongside expanding how much cross-over between the matching articles. The creators have formulated the assessment interaction as referenced under:



**Figure. 2.3.5. Evaluation process of the object detection algorithms**

The creators have involved three calculations for finding the matches between the reference and the result objects. The initial two calculations that is the Bipartite chart coordinating and the Hoover record can be taken on for the division calculations. The third calculation is the clever methodology that is portrayed by the creators that is the multi-object greatest cross-over matching that permits the on-one and many-one affiliations or matching between the reference object and the result object to deal with the under-identification and over-location issues alongside end of the requirement for the limit. Here specific imperatives are laid, one of them is that a given item can be tracked down in just a single matching example. The subsequent imperative is that the matching article ought to have atleast one covering pixel in the result object produced. The fundamental limitation is that the matching articles ought to constantly be ideal. The complete covering region should be expanded. The accuracy, the review rate and the identification precision scores were figured utilizing the assessment techniques. Overall this large number of scores gave corresponding data.



**W. Ouyang and X. Wang[9]**, The thought in the work is to make the visual framework a spectator like the human visual framework which distinguishes the moving items in a characteristic manner. This errand is trying in the machines as the pictures of all articles are expected to get across the retina. The creators in this work have tried the onlookers whether they identify the articles moving when the point of movement is being digressed from the given example of the spiral optic stream . The got data from the 2D picture movement could be utilized for tracking down the heading of movement. The difficult assignment for the spectator is to recognize the moving articles which run over it. To recognize the moving items in the stream field, the point of 2D picture movement is utilized. The creators have led actual investigations to concentrate on the impacts of different seen factors for identification of articles in light of their point of movement. The distinguished the edge point of deviation through a 2D outspread example of movement utilizing which the item can be identified. They have checked the impact of this distinguished limit called as neighborhood versus worldwide signs, the quantity of things referenced in the picture or scene promotion the speed of the moving article. They performed assortment of analyses to see whether the visual framework involves nearby or worldwide signals for object discovery. Object location precision can be accomplished by the visual framework on the off chance that it utilizes the neighborhood prompts. The worldwide signal can impact the item discovery exactness. In another analysis, the creators have tried that while distinguishing the articles in the scene whether the quantity

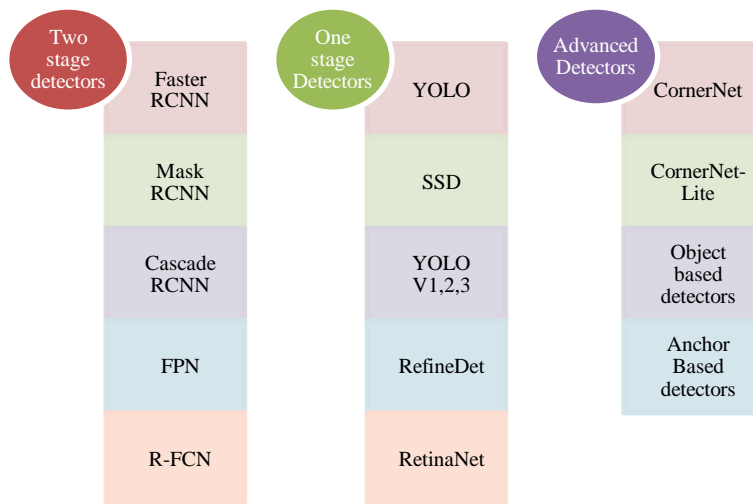
of things in the picture impact the discovery or not. On account of worldwide example, the edges could diminish as the things increment as additional things give more data. In the following two examinations, testing was finished over the speed on the moving item discovery. The edges for identifying a moving item covering the outspread stream field is similarly low in incentive for a deformity design as the visual framework utilizes the worldwide example for recognition. The article location is a lot more straightforward when the 2D picture speed of the boost is expanded.

**Tao Gong, Bin Liu, Qi Chu, Nenghai Yu[10]**, the creators have concocted perform multiple tasks object discovery strategy. This strategy involves the multi-name grouping to acquire improvement the article recognition. The item recognition task embraces the RFCN strategies to address the article identification errands. The multi-mark characterization task is performed by the consideration instrument. The consideration strategy creates the limitation data of the articles. These elements are helpful for playing out the assignment of article discovery. The case level elements as well as the picture level highlights are combined to further develop the article discovery precision. The technique gained needn't bother with any new comment as bouncing box explanations are being utilized for this situation. The assignment of item identification incorporates the errand of distinguishing the example of an article present inside the picture. The R-CNN is being utilized by the creators to play out the item identification is a profound

customary brain network to diminish the calculation time over the example tried. The locale recommendations for this situation are created by determination search or edge box. The Fast r-CNN performs interest pooling to extricate the highlights of the area proposition utilizing the component guide of the picture viable. Afterward, the ROI in the R-FCN by the position delicate ROI pooling to track down the scores of the proposition. The R-FCN applied in the work utilizes completely convolution organization to figure out the common item identification highlight map. Then in the later stage, the R-FCN places its two kin convolution layers after which the PS pooling layer for locale arrangement as well as jumping box relapse on the top. This is treated as a perform various tasks object recognition issue. This work focuses on the multi-name characterization errand and item location task at the same time utilizing just a single profound convolutional network. Gives further developed object discovery precision because of combination of multi name highlights. The structure is named as MONet. The strategy doesn't require additional explanations.

**Payal Mittal, Raman Singh, Akashdeep Sharma [11]**, PC vision has arisen as a field of chances. The Deep learning based object discovery techniques under the PC vision field has a few applications, for example, visual route, object location, sensor based hindrance evasion arranging and some more. The creators in this work have utilized profound learning based object discovery approaches for the

handling of moving items in the pictures those are caught by the robots. This work gives a survey of the profound inclining approaches and headways in the field of item identification. They have worked over the techniques appropriate for object location in the picture informational collection of the low height UAV. The UAV gadgets that are drones are either constrained by human administrator or through installed workstations. These gadgets gather required data or pictures from the low elevation circumstance which now a days has applications in the field of guard, security observation, debacle the board and a few other significant spaces. For the business application, the catching of the UAV pictures and the examination are two primary utilizations of the airborne vehicle. The work referenced here centers around delivering convolution network techniques for object recognition applications. The point is to distinguish the best in class strategy for further developing the item location calculations. The figure underneath shows the whole scientific classification of the customary article recognition calculations alongside the new variants of the item discovery strategies. The figure beneath gives a brief look at the item finders being used:



**Figure. 2.3.6. Taxonomy of Deep learning based Object detection methods**

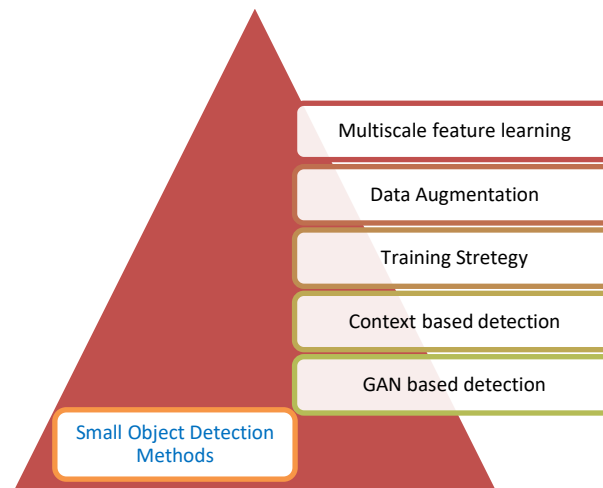
It has been seen that there are sure difficulties in the article recognition techniques pertinent for the UAV informational indexes. The new high level profound learning based object discovery calculations are examined in this work as displayed in the figure above. The high level discovery techniques are more reasonable for the identification in the event of the low height based flying items. For these techniques to work for object discovery, the component tensor utilized for pooling should be standardized before connection. To acquire the better article highlights, enough pixels ought to be used to identify little items in the picture and those ought to address the enough subtleties of the article.

**Kang Tong, Yiquan Wu, Fei Zhou[12],** The PC Vision is the huge field which focuses on issues that can computerize the errand of item discovery in numerous areas. The creators in this work have examined about the little article recognition techniques and approaches in view of profound learning. The techniques incorporate multiscale include learning,

setting based discovery, information increase, GAN-based location alongside preparing procedure. The little article location techniques work on the discovery of examples of little objects of the specific class in the information pictures and recordings. The little item discovery is material for undertakings including case division, object following, activity acknowledgment, picture subtitling, and so on. In the more extensive degree it has been generally utilized in the space of advanced mechanics vision, astute transportation, independent driving, reconnaissance and numerous others. The little items can be characterized as far as its actual estimations or the quantity of pixels used to address the articles. The calculations are seen performing better in the event of huge items, yet in the event of little articles there are sure difficulties. The difficulties remember the absence of appearance of the item for the picture because of comparable foundations, the precise restrictions of the article isn't basic, the information expected for little article discovery is exceptionally restricted. In this work, a deliberate survey over little

item identification techniques is portrayed, the presentation measures for little item discovery is examined and the future headings are referenced nearby.

The techniques for the little item location depicted in this work is portrayed in the graph beneath:



**Figure. 2.3.7. Small Object Detection Approaches**

The multiscale highlight learning focuses on tending to not many learning issues over little items including single element map, pyramidal combination, highlight pyramid age and organization age, highlight combination, multiscale combination and featurized picture pyramids. The majority of the finders depend on using the featurized picture pyramids. The Data Augmentation implies bothering the info picture through couple of changes including the essential flipping, revolution, scaling, trimming, and so forth. The justification for applying these is to produce extra examples for additional location cycle. The information increase is appropriate for preparing and testing. For taking care of the intricacy, another strategy is proposed named scale standardization for picture pyramids which

back spreads the inclinations of the articles occurrences for compelling preparation of the techniques. During the preparation, ground truth boxes and recommendations which fall in a particular size range for a particular goal.

**Conclusion:**

The multiscale highlight learning focuses on tending to not many learning issues over little items including single element map, pyramidal combination, highlight pyramid age and organization age, highlight combination, multiscale combination and featurized picture pyramids. The majority of the finders depend on using the featurized picture pyramids. The Data Augmentation implies bothering the info picture through couple of changes including the essential flipping, revolution, scaling, trimming, and so forth. The justification for applying these is to produce extra examples for additional location cycle. The information increase is appropriate for preparing and testing. For taking care of the intricacy, another strategy is proposed named scale standardization for picture pyramids which back spreads the inclinations of the articles occurrences for compelling preparation of the techniques. During the preparation, ground truth boxes and recommendations which fall in a particular size range for a particular goal.

**References:**

- [1]. D. Moshou, C. Bravo, R. Oberti, J. West, L. Bodria, A. McCartney, H. Ramon, "Plant disease detection based on data fusion of hyper-spectral and multi-spectral fluorescence imaging using Kohonen maps", *Real-Time Imaging*, Volume 11, Issue 2, 2005, Pages 75-83, ISSN 1077-2014, doi:10.1016/j.rti.2005.03.003.
- [2]. W. Ouyang and X. Wang, "A discriminative deep model for pedestrian detection with occlusion handling," *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 3258-3265, doi: 10.1109/CVPR.2012.6248062.
- [3]. Richa Singh, Mayank Vatsa, Afzel Noore, "Hierarchical fusion of multi-spectral face images for improved recognition performance", *Information Fusion*, Volume 9, Issue 2, 2008, Pages 200-210, ISSN 1566-2535, doi:10.1016/j.inffus.2006.06.002.
- [4]. Leitner, R., Biasio, M., De Arnold, T., Dinh, C.V., Loog, M., Duin, R.P.W., "Multi-spectral video endoscopy system for the detection of cancerous tissue", *Pattern Recognit. Lett.* **34**, 85–93 (2013)
- [5]. Volker Walter, "Object-based classification of remote sensing data for change detection", *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 58, Issues 3–4, 2004, Pages 225-238, ISSN 0924-2716, doi:10.1016/j.isprsjprs.2003.09.007.
- [6]. Ian Fasel, Bret Fortenberry, Javier Movellan, "A generative framework for real time object detection and classification", *Computer Vision and Image Understanding*, Volume 98, Issue 1, 2005, Pages 182-210, ISSN 1077-3142, doi:10.1016/j.cviu.2004.07.014.
- [7]. asaaki Shibata, Yuichiro Yasuda, Masahide Ito, "Moving Object Detection for Active Camera based on Optical Flow Distortion", *IFAC Proceedings Volumes*, Volume 41, Issue 2, 2008, Pages 14720-14725, ISSN 1474-6670, ISBN 9783902661005, doi:10.3182/20080706-5-KR-1001.02492.
- [8]. Bahadır Özdemir, Selim Aksoy, Sandra Eckert, Martino Pesaresi, Daniele Ehrlich, "Performance measures for object detection evaluation, *Pattern Recognition Letters*, Volume 31, Issue 10, 2010, Pages 1128-1137, ISSN 0167-8655, doi:10.1016/j.patrec.2009.10.016.

- [9]. Constance S. Royden, Erin M. Connors," The detection of moving objects by moving observers", *Vision Research*, Volume 50, Issue 11, 2010, Pages 1014-1024, ISSN 0042-6989, doi:10.1016/j.visres.2010.03.008.
- [10]. Tao Gong, Bin Liu, Qi Chu, Nenghai Yu,"Using multi-label classification to improve object detection", *Neurocomputing*, Volume 370, 2019, Pages 174-185, ISSN 0925-2312, doi:10.1016/j.neucom.2019.08.089.
- [11]. Payal Mittal, Raman Singh, Akashdeep Sharma," Deep learning-based object detection in low-altitude UAV datasets: A survey", *Image and Vision Computing*, Volume 104, 2020, 104046, ISSN 0262-8856, doi:10.1016/j.imavis.2020.104046.
- [12]. Kang Tong, Yiquan Wu, Fei Zhou,"Recent advances in small object detection based on deep learning: A review", *Image and Vision Computing*, Volume 97, 2020, 103910, ISSN 0262-8856, doi:10.1016/j.imavis.2020.103910.
- [13] Ajani, S.N., Amdani, S.Y. (2021). Agent-Based Path Prediction Strategy (ABPP) for Navigation Over Dynamic Environment. In: Muthukumar, P., Sarkar, D.K., De, D., De, C.K. (eds) *Innovations in Sustainable Energy and Technology. Advances in Sustainability Science and Technology*. Springer, Singapore. [https://doi.org/10.1007/978-981-16-1119-3\\_16](https://doi.org/10.1007/978-981-16-1119-3_16)
- [14] S. N. Ajani and S. Y. Amdani, "Probabilistic path planning using current obstacle position in static environment," *2nd International Conference on Data, Engineering and Applications (IDEA)*, 2020, pp. 1-6, doi: 10.1109/IDEA49133.2020.9170727.
- [15] S. Ajani and M. Wanjari, "An Efficient Approach for Clustering Uncertain Data Mining Based on Hash Indexing and Voronoi Clustering," *2013 5th International Conference and Computational Intelligence and Communication Networks*, 2013, pp. 486-490, doi: 10.1109/CICN.2013.106.
- [16] Ajani, S. and Amdani, S.Y., 2020. Dynamic Path Planning Approaches based on Artificial Intelligence and Machine Learning. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 11(3), pp.2084-2098.
- [17] Ajani, S.N. and Wanjari, M., 2013. An approach for clustering uncertain data objects: A survey.[J]. *International Journal of Advanced Research in Computer Engineering & Technology*, 2, p.6.
- [18] Ajani, S., Amdani, S.Y. and Asole, S.S., 2021. Probabilistic Path Planning over Dynamic environment using Improved Dynamic RRT\* Algorithm. *Design Engineering*, pp.6613-6621.